



SAPIENZA  
UNIVERSITÀ DI ROMA

ISSN 2385-2755

DiSSE Working papers  
[online]

**WORKING PAPERS SERIES**  
**DIPARTIMENTO DI**  
**SCIENZE SOCIALI ED ECONOMICHE**

# The inclusive synthetic control method

**Roberta Di Stefano, Giovanni Mellace**



**N. 21/2020**

SAPIENZA - UNIVERSITY OF ROME

P.le Aldo Moro n.5 – 00185 Roma T(+39) 0649910563

CF80209930587 – P.IVA 02133771002

# The inclusive synthetic control method

Roberta Di Stefano\*      Giovanni Mellace†

November 24, 2020

## Abstract

The Synthetic Control Method (SCM) estimates the causal effect of a policy intervention in a panel data setting with only a few treated units and control units. The treated outcome in the absence of the intervention is recovered by a weighted average of the control units. The latter cannot be affected by the intervention, neither directly nor indirectly. We introduce the inclusive synthetic control method (iSCM), a novel and intuitive synthetic control modification that allows including units potentially affected directly or indirectly by an intervention in the donor pool. Our method is well suited for applications with multiple treated units where including treated units in the donor pool substantially improves the pre-intervention fit and/or for applications where some of the units in the donor pool might be affected by spillover effects. Our iSCM is very easy to implement, and any synthetic control type estimation and inference procedure can be used. Finally, as an illustrative empirical example, we re-estimate the causal effect of German reunification on GDP per capita allowing for spillover effects from West Germany to Austria.

**Keywords:** Synthetic Control Method, spillover effects, causal inference.

**JEL classification:** C21, C23, C31, C33.

---

\*Department of Methods and Models for Economics, Territory and Finance Sapienza University of Rome Via del Castro Laurenziano 9, Rome, 00161 roberta.distefano@uniroma1.it

†Department of Business and Economics, Campusvej 55, 5230 Odense M, Denmark, giome@sam.sdu.dk

# 1 Introduction

The synthetic control method (SCM) introduced by Abadie and Gardeazabal (2003) and further developed in Abadie et al. (2010) and Abadie et al. (2015) allows estimating the causal effect of a policy intervention in setting where only a few treated and control units are observed over a long time period. The idea behind this method is to create a linear combination of control units that mimic what would have happened to the treated units in the absence of the intervention. The weights given to each control unit are chosen to minimize the distance in the pre-intervention outcomes of the treated and the synthetic control. The causal effect of the intervention is estimated as the difference between the observed outcome of treated and the one of the synthetic control in the post-intervention period.

One of the key assumptions of SCM is that only units that are not affected by the interventions are included in the control group, often referred to as donor pool. This might be problematic in at least two scenarios: i) some of the treated units need to be included in the donor pool for the treated to be in the convex hull of the control units; ii) some of the control units in the donor pool are affected by the intervention indirectly. As a motivating example, consider the German reunification study of Abadie et al. (2015) and Abadie and L'Hour (2019), as the authors point out it is possible that German reunification had spillover effects on a neighbor country like Austria. As Austria receives a high weight (42%) in their study, such a spillover effect, if large, would introduce a large bias. Given that Austria plays an important role in constructing “synthetic West Germany”, excluding it from the donor pool is likely to induce violations of the SCM assumptions as it is less plausible that West Germany is in the convex hull of the other control units.

Our main contribution is to introduce the inclusive synthetic control method (iSCM), a novel procedure that allows us to eliminate post-intervention effects from control units and safely include them in the donor pool. Our procedure does not require to modify the original synthetic control estimator, and all the new recent methods can be used instead. The main additional assumptions required are that the number of “potentially affected” units is known and that the standard SCM assumptions would hold if there were no post-intervention effects for those units.

Although iSCM only requires the existence of at least one “pure control” unit, we expect that the quality of our estimator deteriorates if the number of “potentially affected” units increases. Thus, it is advisable to impose assumptions that limit this number. This is similar to what is done in the literature on spillover effects, where it is often assumed that interactions between units are only possible in the same group but not between different groups (Cerqua and Pellegrini 2017, Forastiere et al. 2016, Huber and Steinmayr 2019, Vazquez-Bare 2017). This is also the case for contributions to this literature based on synthetic control. Grossi et al. (2020) propose to reduce the donor pool to only units not affected by spillovers. They estimate the effect for the treated unit using a standard SCM with the restricted donor pool and the spillover effects comparing units affected by spillover and the restricted donor pool. Their method is very effective in applications where the restricted donor pool is sufficient to construct a “good” synthetic control. However, in a setting where the units affected by spillover need to be included in the donor pool, as in the German reunification example, their method would likely produce biased results. Cao and Dowd (2019) provide a different identification strategy imposing a linear spillover structure restricting effect heterogeneity. In contrast, our approach does not impose any assumptions on the spillover effect heterogeneity. The rest of the paper is organized as follows: Section 2 reviews the literature; Section 3 introduces our iSCM; Section 4 proposes possible inference procedures; Section 5 presents the results of empirical applications; and Section 6 concludes.

## 2 Literature review

SCM is receiving increasing attention in the literature. Athey and Imbens (2017) argue that SCM is “...the most important innovation in the policy evaluation literature in the last 15 years”. Several contributions have been made to improve upon the original method (see Abadie (2020) for a recent review of the literature) in several dimensions. One strand of the literature focuses on bias reduction due to imbalance in observed characteristics. Abadie and L’Hour (2019) propose a bias reduction procedure based on introducing a penalty term that reduces pairwise matching discrepancies between the characteristics of the treated and each of the

control units and helps to avoid interpolation bias. Botosaru and Ferman (2019) discuss implications of not having perfect covariate balance and provide alternative assumptions under which SCM can still be used. Kellogg et al. (2020) propose a model averaging method called “matching synthetic control estimator” that is a convex combination of the synthetic control and matching estimators. Their procedure gives weight to the synthetic control estimator that are proportional to the risk of having extrapolation bias.

Another strand of the literature focuses on problems related to have an imperfect fit in the pre-treatment period. Ferman and Pinto (2019) analyze the properties of SCM when the pre-treatment fit is imperfect. Similarly, Ben-Michael et al. (2020) discuss potential problems with SCM and propose an outcome model to estimate the bias. They also consider staggered adoption setting (Ben-Michael et al. 2019). Doudchenko and Imbens (2017) allow a better pre-intervention fit, proposing a generalization of the synthetic control, relaxing weight-constraints, i.e., allowing weights to be negative, and their sum to be different to one, and adding a time-constant intercept.

Finally, several contributions focus on generalizing the method and compare it to alternative approaches. Gobillon and Magnac (2016) compare linear factor models and synthetic controls. Xu (2017) proposes a generalization to unify synthetic control with linear fixed-effects models. Amjad et al. (2018) propose a procedure based on de-noising the outcomes and imputing the missing values. Arkhangelsky et al. (2019) propose a new synthetic control as a weighted regression estimator with time fixed effects. Mellace and Pasquini (2019) show how to use SCM to estimate how much of the total effect of intervention goes through observed intermediate outcomes (causal channels). Athey et al. (2020) use matrix completion techniques to derive a new method that include synthetic control as a special case. There are contributes that exploit the connection between SCM and other approaches.

### **3 The inclusive synthetic control method**

Without loss of generality assume we are interested in the effect of an intervention, implemented at time  $T$ , on an outcome  $Y$  of one treated unit. We will refer to this

unit as the “main treated”. We assume to observe  $J$  units ordered such that unit 1 is the “main” treated, units 2 to  $m \leq J - 1$  (potentially “affected” hereafter) are either other treated units that we would like to include in the donor pool or control units that might be affected by spillover effects from the main treated, and units  $m + 1$  through  $J$  are “pure” control units that are not affected by the intervention at all.

We define the potential outcome (see, e.g., Rubin 1974)  $Y_{1t}^I$  as the outcome that the main treated unit would obtain under the intervention at time  $t$ . With a little abuse of notation and depending on the specific application,  $Y_{jt}^S, j = 2, \dots, m$  represent the potential outcomes at time  $t$  in the presence of the intervention received by either the other treated units or by the units potentially affected by spillover effects. Finally, we define as  $Y_{jt}^N, j = 1, \dots, J$  the potential outcome in the absence of the intervention. We denote the number of pre-intervention periods as  $T_0$  and we define the following two binary indicators

$$D_{jt} = \begin{cases} 1 & \text{if } j = 1 \text{ and } t > T_0, \\ 0 & \text{otherwise.} \end{cases}$$

$$S_{jt} = \begin{cases} 1 & \text{if } j = 2, \dots, m \text{ and } t > T_0, \\ 0 & \text{otherwise.} \end{cases}$$

These binary indicators are used to select the “main” treated and the units that are potentially affected by the intervention, respectively, in the post-intervention period.

Assuming no anticipation effects in the pre-treatment period and that the standard stable unit treatment value assumption (SUTVA) holds (partially in the case of spillover effects), we can relate the observed and the potential outcome by the following observational rule

$$Y_{jt} = Y_{jt}^N(1 - D_{jt})(1 - S_{jt}) + Y_{jt}^I D_{jt} + Y_{jt}^S S_{jt}.$$

This implies that in the pre-intervention period,  $Y_{jt} = Y_{jt}^N$  for all units, while

in the post-intervention period,  $Y_{jt}=Y_{jt}^N$  for the “pure” control units;  $Y_{1t}=Y_{1t}^I$  for “main” treated and  $Y_{jt}=Y_{jt}^S$  for the other potentially affected units.

Our parameters of interest are the effect of the intervention for the main treated at time  $t > T_0$ , denoted by  $\theta_{1t}$ , and the effects on the other potentially affected units denoted by  $\gamma_{jt}, j = 2, \dots, m, t > T_0$ , defined as

$$\theta_{1t} = Y_{1t}^I - Y_{1t}^N, \quad t > T_0$$

and

$$\gamma_{jt} = Y_{jt}^S - Y_{jt}^N, \quad j = 2, \dots, m, \quad t > T_0.$$

To identify these parameters, we need to recover  $Y_{1t}^N$  and  $Y_{jt}^N$  for  $j = 2, \dots, m$  in the post-treatment period. If, hypothetically, one used the standard SCM as described in Abadie et al. 2010 and included the potentially affected units in the donor pool, the resulting estimate of the counterfactual potential outcome of the main treated in the absence of the intervention would be

$$\widehat{Y}_{1t}^N = \sum_{j=2}^J w_j^* Y_{jt},$$

where the  $(J \times 1)$  vector of weights  $W^* = (w_2^*, \dots, w_J^*)'$  is chosen to minimize the distance between the treated and the other units in pre-intervention characteristics. Thus, the effect on the main treated would be estimated as

$$\widehat{\theta}_{1t} = Y_{1t} - \sum_{j=2}^J w_j^* Y_{jt}.$$

As units 2 to  $m$  are potentially affected by the intervention, their post intervention outcomes are given by

$$Y_{jt} = Y_{jt}^N + \gamma_{jt}, \quad j = 2, \dots, m.$$

Our first assumption is that if units 2 to  $m$  were not affected by the intervention, the standard SMC would work, formally

**Assumption 1:** There exists a set of weights  $w_j^*, j = 2, \dots, J$  such that

$$Y_{1t}^N = \sum_{j=2}^J w_j^* Y_{jt}^N.$$

In other words, Assumption 1 ensures that if units 2 to  $J$  were not affected by the intervention, the potential outcome of the treated unit would be in their convex hull.

**Lemma 1:** Under Assumption 1

$$\hat{\theta}_{1t} = \theta_{1t} - \sum_{j=2}^m w_j^* \gamma_{jt} \quad (1)$$

**Proof of Lemma 1:** Under Assumption 1, using the observational rule, we have

$$\begin{aligned} \hat{Y}_{1t}^N &= \sum_{j=2}^J w_j^* Y_{jt} \\ &= \sum_{j=m+1}^J w_j^* Y_{jt}^N + \sum_{j=2}^m w_j^* (Y_{jt}^N + \gamma_{jt}) \\ &= \sum_{j=2}^J w_j^* Y_{jt}^N + \sum_{j=2}^m w_j^* \gamma_{jt} \\ &= Y_{1t}^N + \sum_{j=2}^m w_j^* \gamma_{jt} \end{aligned}$$

This immediately implies that

$$\hat{\theta}_{1t} = \theta_{1t} - \sum_{j=2}^m w_j^* \gamma_{jt}.$$

□

Lemma 1 shows how the presence of post interventions effects affects the standard SCM under assumption 1.

**Remark:** It is important to notice that for each unit  $j = 2, \dots, m$  if either  $\gamma_{jt}$  or  $w_j^*$  is zero that unit does not induce “bias” in  $\hat{\theta}_{1t}$ . This implies that units



that receive a low estimated weight need to have a large effect to induce bias in  $\widehat{\theta}_{1t}$ . For this reason, units that receive a low weight using the standard SCM can be relatively safely treated as pure controls when estimating  $\theta_{1t}$  in empirical applications.

Consider a generic potentially affected unit  $i = 2, \dots, m$ . Using a standard SCM to estimate  $Y_{it}^N$  also including the main treated (unit 1) and all other  $m - 1$  affected units in the donor pool would require finding a vector of weights  $L^{*i}$ , such that

$$\widehat{Y}_{it}^N = \sum_{j \neq i} l_j^{*i} Y_{jt}.$$

Let  $\mathcal{J} = \{1, \dots, J\}$ , we assume that for units 2 to  $m$  without the effect of the intervention on the main treated and the other potentially affected units the standard SMC would work, formally

**Assumption 2:** There exists a set of weights  $l_j^{*i}$ ,  $j \in \mathcal{J} \setminus \{i\}$ , such that

$$Y_{it}^N = \sum_{j \in \mathcal{J} \setminus \{i\}} l_j^{*i} Y_{jt}^N, \quad \forall i = 2, \dots, m.$$

**Lemma 2:** Under Assumption 2

$$\widehat{\gamma}_{it} = \gamma_{it} - \sum_{j \in M \setminus \{i\}} l_j^{*i} \gamma_{jt} - l_1^{*i} \theta_{1t}. \quad (2)$$

**Proof of Lemma 2:** Under Assumption 2, we have

$$\widehat{Y}_{it}^N = Y_{it}^N + \sum_{j \in M \setminus \{i\}} l_j^{*i} \gamma_{jt} + l_1^{*i} \theta_{1t},$$

with  $M = \{2, \dots, m\}$ . It follows

$$\widehat{\gamma}_{it} = \gamma_{it} - \sum_{j \in M \setminus \{i\}} l_j^{*i} \gamma_{jt} - l_1^{*i} \theta_{1t}.$$

□

Combining the results of Lemma 1 and Lemma 2, we obtain the following system of equations:

$$\begin{aligned}
\widehat{\theta}_{1t} &= \theta_{1t} - \sum_{j \in M} w_j^* \gamma_{jt} \\
\widehat{\gamma}_{2t} &= \gamma_{2t} - \sum_{j \in M \setminus \{2\}} l_j^{*2} \gamma_{jt} - l_1^{*2} \theta_{1t} \\
\widehat{\gamma}_{3t} &= \gamma_{3t} - \sum_{j \in M \setminus \{3\}} l_j^{*3} \gamma_{jt} - l_1^{*3} \theta_{1t} \\
&\dots \\
\widehat{\gamma}_{mt} &= \gamma_{jt} - \sum_{j \in M \setminus \{m\}} l_j^{*m} \gamma_{jt} - l_1^{*m} \theta_{1t}
\end{aligned}$$

After some simple manipulations we obtain:

$$\begin{aligned}
\widehat{\theta}_{1t} &= \theta_{1t} - w_2^* \gamma_{2t} - w_3^* \gamma_{3t} - \dots - w_m^* \gamma_{mt} \\
\widehat{\gamma}_{2t} &= -l_1^{*2} \theta_{1t} + \gamma_{2t} - l_3^{*2} \gamma_{3t} - \dots - l_m^{*2} \gamma_{mt} \\
\widehat{\gamma}_{3t} &= -l_1^{*3} \theta_{1t} - l_2^{*3} \gamma_{2t} + \gamma_{3t} - \dots - l_m^{*3} \gamma_{mt} \\
&\dots \\
\widehat{\gamma}_{mt} &= -l_1^{*m} \theta_{1t} - l_2^{*m} \gamma_{2t} - l_3^{*m} \gamma_{3t} - \dots + \gamma_{mt}
\end{aligned}$$

This is a system of  $m$  equations with  $m$  unknowns, i.e., the treatment effect on the main treated and the  $m - 1$  effects on the potentially affected units.

We can write this system in matrix form, denoting by  $\vartheta_t$  the  $(m \times 1)$  vector of unknown parameters (our effects of interest), by  $\Omega$  the  $(m \times m)$  matrix of known quantities (our estimated weights) that has ones on the main diagonal and by  $\beta_t$  the  $(m \times 1)$  vector of known quantities (biased estimated effects), as

$$\beta_t = \begin{pmatrix} \widehat{\theta}_{1t} \\ \widehat{\gamma}_{2t} \\ \widehat{\gamma}_{3t} \\ \vdots \\ \widehat{\gamma}_{mt} \end{pmatrix} \quad \Omega = \begin{pmatrix} 1 & -w_2^* & -w_3^* & \dots & -w_m^* \\ -l_1^{*2} & 1 & -l_3^{*2} & \dots & -l_m^{*2} \\ -l_1^{*3} & -l_2^{*3} & 1 & \dots & -l_m^{*3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -l_1^{*m} & -l_2^{*m} & -l_3^{*m} & \dots & 1 \end{pmatrix} \quad \vartheta_t = \begin{pmatrix} \theta_{1t} \\ \gamma_{2t} \\ \gamma_{3t} \\ \vdots \\ \gamma_{mt} \end{pmatrix} \quad (3)$$

We now assume that  $\Omega$  is invertible, namely

**Assumption 3:**  $\Omega$  is non-singular.

It is easy to show that  $\Omega$  is always invertible, if  $m \leq J - 1$ , except for the extreme cases where two units give weight 1 to each other and/or every single weight associated with the pure control units is zero (see Appendix A).

We now state our main result in the following theorem.

**Theorem 1:** Under Assumption 3, we have

$$\vartheta_t = \Omega^{-1}\beta_t.$$

**Proof of Theorem 1:** The result immediately follows from equation 3 using the fact that  $\Omega$  is invertible.  $\square$

The result in Theorem 1 can be readily used to identify our effects of interest by simply applying Cramer’s rule:

$$\vartheta_{jt} = \frac{\det(\Omega_{j,t})}{\det(\Omega)} \quad j = 1, \dots, m.$$

where  $\Omega_{j,t}$  is the matrix obtained by replacing the  $j$ -th column of  $\Omega$  by the vector  $\beta_t$ .

The expression above makes it very easy to construct estimators of our parameters of interest that only require very basic linear algebra operations together with any SCM-type estimator for the weight matrix  $\Omega$  and the vector  $\beta_t$ .

To further illustrate our results, it is useful to consider the special case where, together with the “main treated unit”, only one additional unit is potentially affected by the intervention ( $m = 1$ ).

In this case our system of equations simplifies to:

$$\begin{cases} \hat{\theta}_t = \theta_t - w_2^* \gamma_t \\ \hat{\gamma}_t = -l_1^* \theta_t + \gamma_t \end{cases}$$

Therefore, we have

$$\beta_t = \begin{pmatrix} \widehat{\theta}_t \\ \widehat{\gamma}_t \end{pmatrix}, \quad \Omega = \begin{pmatrix} 1 & -w_2^* \\ -l_1^* & 1 \end{pmatrix}, \quad \vartheta_t = \begin{pmatrix} \theta_t \\ \gamma_t \end{pmatrix}.$$

To derive expressions for our parameters of interest we need to find  $\det(\Omega)$ ,  $\det(\Omega_{1,t})$  and  $\det(\Omega_{2,t})$ , which are given by

$$\det(\Omega) = \begin{vmatrix} 1 & -w_2^* \\ -l_1^* & 1 \end{vmatrix} = 1 - w_2^* l_1^*,$$

$$\det(\Omega_{1,t}) = \begin{vmatrix} \widehat{\theta}_t & -w_2^* \\ \widehat{\gamma}_t & 1 \end{vmatrix} = \widehat{\theta}_t + w_2^* \widehat{\gamma}_t,$$

$$\det(\Omega_{2,t}) = \begin{vmatrix} 1 & \widehat{\theta}_t \\ -l_1^* & \widehat{\gamma}_t \end{vmatrix} = \widehat{\gamma}_t + l_1^* \widehat{\theta}_t.$$

Following Cramer's rule we obtain

$$\theta_t = \frac{\widehat{\theta}_t + w_2^* \widehat{\gamma}_t}{1 - w_2^* l_1^*}$$

$$\gamma_t = \frac{\widehat{\gamma}_t + l_1^* \widehat{\theta}_t}{1 - w_2^* l_1^*}$$

In this case, it is easy to see that  $\det(\Omega)$  is always different from zero, except if  $w_2^* = l_1^* = 1$ . Thus, our parameters of interest are always identified unless the main treated gives weight 1 to the other affected unit which in turns gives weight 1 to the main treated. This would be the case, for example, if there are no “pure control” units.

## 4 Inference

Dealing with only a few units, makes inference for synthetic control based methods, like ours, complicated. We can, however, easily adapt existing methods to our setting. The most popular choice is to implement permutation tests. Abadie et al. (2010) and Abadie et al. (2015) propose placebo tests in time, i.e., reassigning the intervention artificially before its real implementation and placebo tests in space, i.e., reassigning the intervention artificially for units in the control group. The latter approach is often preferred because of possible shocks that might have occurred in the past affecting units differently. In space placebo tests measure the statistical significance of the effect through the ratio between the root mean squared prediction errors (RMSPE) in the post-treatment period and in the pre-treatment period. The RMSPE measures the lack of fit between the observed outcome and its synthetic control. In our framework the presence of units affected by intervention in the donor pool, requires a small modification in the way we compute the post-intervention RMSPE. We suggest computing the post-intervention RMSPE by subtracting from the outcomes of each affected unit (excluding the one for which we are estimating the effect) the respective effect estimated with iSCM. For the main treated units the modified RMSPE ratio becomes

$$r_1 = \frac{\left(\frac{1}{T-T_0} \sum_{t=T_0+1}^T (Y_{1t} - (\hat{Y}_{1t}^N - \sum_{j=2}^m w_j^* \gamma_{jt}))^2\right)^{1/2}}{\left(\frac{1}{T_0} \sum_{t=1}^{T_0} (Y_{1t} - \hat{Y}_{1t}^N)^2\right)^{1/2}},$$

while for the other potentially affected units we have

$$r_j = \frac{\left(\frac{1}{T-T_0} \sum_{t=T_0+1}^T (Y_{jt} - (\hat{Y}_{jt}^N - \sum_{j \in M \setminus \{i\}} l_j^{*i} \gamma_{jt} - l_1^{*i} \theta_{1t}))^2\right)^{1/2}}{\left(\frac{1}{T_0} \sum_{t=1}^{T_0} (Y_{jt} - \hat{Y}_{jt}^N)^2\right)^{1/2}}, j = 2, \dots, m.$$

This idea can be easily applied to other inference procedures available in the literature (see, e.g., Cao and Dowd 2019; Chernozhukov et al. 2018; Firpo and Possebom 2018; Gobillon and Magnac 2016; Li 2019).

## 5 Empirical example

In this section, we use iSCM to estimate the effect of German reunification on West Germany’s per capita GDP. In this application, one of the control units (Austria) in the donor pool is potentially affected indirectly by the treatment<sup>1</sup>. As discussed in Abadie et al. (2015) and Abadie and L’Hour (2019), German reunification could have had negative spillover effects on Austria’s economic growth because West Germany diverted demand and investment from Austria to East Germany. This would imply that the big negative effect that they found is likely to be an upper bound of the true effect. As it is arguably important to include Austria in the donor pool, our method is very well suited for this empirical application. In October 1990, less than a year after the fall of the Berlin wall on November 1989, the German Democratic Republic (“East Germany”) and the Federal Republic of Germany (“West Germany”) were officially reunified. The differences between the two economies were large. In 1989 the GDP per capita of West Germany was about three times higher than that of East Germany (Schinasi et al. 1990). German reunification, defined as one of the most important historical milestones of European history after 1945, most likely affected not only the German economy but also other countries. In particular, Austria has had tight links with Germany historically, also because the two countries share the same language and, to a great extent, a common history. In 1938, Austria was annexed by the Third Reich that benefited from its raw materials and labor to complete the German rearmament. In 1945, Austria was separated from Germany. However, the economic cooperation between Austria and West Germany continued during the Cold War.

We use the same specification as in Abadie et al. (2015) to estimate the synthetic for West Germany. In order to find the weights to assign to each covariate, they split the pre-treatment period in a training period (1971–80) and in a validation period (1981–90). The weights are then selected by minimizing the out-of-sample error in the validation period. For Austria, we cannot safely use this procedure. As described in Gehler and Graf (2018), in 1980, right before the sample split cut-off, Austria provided several loans to East Germany, and in return,

---

<sup>1</sup>Given that other European countries receive very little weights the impact of potential spillover effects on those countries would arguable be negligible as shown by Lemma 1.

its nationalized industries received large-scale orders. This most likely fostered Austria's exports and contributed to jobs creation in its industries. Thus, the sample split might catch the effect of this economic shock. This is corroborated by the fact that using the same specification as in Abadie et al. (2015) also for Austria leads to a bad pre-treatment fit. For this reason, we decided to follow Abadie et al. (2010) in choosing the covariates weights for synthetic Austria, which are selected such that the mean squared prediction error of the outcome variable is minimized for the entire pre-treatment period. We use country-level panel data that cover the period 1960-2003, with our post-intervention period starting in 1990. Except for Austria, the "pure control" countries in the donor pool include 15 other OECD countries: Australia, Belgium, Denmark, France, Greece, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Switzerland, the United Kingdom, and the United States. The outcome variable is the real per capita GDP at Purchasing Power Parity (PPP) measured in 2002 USD. The pre-intervention covariates include: per capita GDP, inflation rate, industry share of value added, investment rate, schooling, and a measure of trade openness.

Applying iSCM requires the following steps:

1. After constructing Synthetic West Germany using the entire donor pool (including Austria), estimate the bias treatment effect  $\hat{\theta}_t$  and the weight assigned to Austria  $\hat{w}_A$ .
2. After constructing Synthetic Austria including West Germany in the donor pool, estimate the bias spillover effect  $\hat{\gamma}_t$  and the weight assigned to West Germany  $\hat{l}_{WG}$ .
3. Estimate the unbiased treatment effect on West Germany as  $\frac{\hat{\theta}_t + \hat{w}_A \hat{\gamma}_t}{1 - \hat{w}_A \hat{l}_{WG}}$ .
4. Estimate the unbiased spillover effect on Austria as  $\frac{\hat{\gamma}_t + \hat{l}_{WG} \hat{\theta}_t}{1 - \hat{w}_A \hat{l}_{WG}}$ .

Step 1 allows us also to judge whether West Germany gives enough weight to Austria to induce a non-negligible bias. Similarly, after step 2 we can tell whether Austria gives enough weight to the West Germany to have a large bias in the estimation of the spillover effect. Table 1 shows the estimated weights of synthetic West Germany in the second column, synthetic Austria in the third column, and

synthetic West Germany excluding Austria from the donor pool in the fourth. We can observe that West Germany gives Austria the highest weight (42%) and that Austria gives also the the highest weight (33%) to West Germany. When we exclude Austria from the donor pool, West Germany becomes a weighted average of the USA, Netherlands, Japan, and Switzerland; however, the pretreatment fit is way worse, as shown in Figures 1 and 2 below. After steps 1 and 2 are implemented, we can check whether Assumption 3, i.e., the non-singularity of the matrix  $\Omega$ , holds.  $\Omega$  in this example is given by

$$\Omega = \begin{pmatrix} 1 & -0.42 \\ -0.33 & 1 \end{pmatrix}$$

As  $\det(\Omega) = 0.86$ , Assumption 3 holds in this application and we can now proceed to steps 3 and 4. Specifically, we need to find  $\det(\Omega_{WG,t})$  and  $\det(\Omega_{A,t})$  for each period, where  $\Omega_{WG,t}$  and  $\Omega_{A,t}$  are matrices obtained by replacing in  $\Omega$  the vector of estimated effects  $\beta_t$  in the first column for West Germany and in the second column for Austria, namely:

$$\Omega_{WG,t} = \begin{pmatrix} \hat{\theta}_t & -0.42 \\ \hat{\gamma}_t & 1 \end{pmatrix} \text{ and } \Omega_{A,t} = \begin{pmatrix} 1 & \hat{\theta}_t \\ -0.33 & \hat{\gamma}_t \end{pmatrix}.$$

The treatment and spillover effects for each period are given by  $\frac{\det(\Omega_{WG,t})}{\det(\Omega)}$  and  $\frac{\det(\Omega_{A,t})}{\det(\Omega)}$ , respectively. The results are shown in Figure 1, where we can see the per capita GDP trajectory of West Germany, its synthetic counterpart in the standard synthetic control version (including spillover effect), in the “restricted” synthetic control version (excluding Austria from the donor pool), and in the inclusive synthetic control version (not including the spillover effect), in the 1960–2003 period. We can see that the standard and inclusive synthetic version of West Germany in the pre-reunification period reproduce almost perfectly West Germany per capita GDP, while excluding Austria substantially deteriorates the pre-reunification fit. This confirms the importance of including Austria in the donor pool. Abadie et al. (2015) find a negative effect of the reunification on West Germany per capita GDP that was reduced by approximately 7.67% per year on average with respect to the 1990 baseline level. Our iSCM results are not very different from the one of



Abadie et al. (2015) and confirm their expectation about the potential direction of the bias, which implies an even more negative effect of reunification. The difference between the trends in per capita GDP between iSCM and SCM is generally small, as better shown in Figure 2 and in Table 2. Our iSCM estimate implies a negative effect that is up to 1.50% larger than the one estimated with a standard SCM. Figure 3 and Table 3 show the gap between iSCM Austria and Austria and leads to similar conclusions. Austria's per capita GDP in 1997-1998 and 2001 is about 700 USD per year less than it would have been in the absence of reunification. Finally, Figure 4 shows the ratios between the RMSPEs in the post- and pre-reunification of West Germany and the donor pool. We can observe that the value for West Germany is very high and the largest compared to the other countries in the donor pool.

Table 1: Synthetic control weights for West Germany and Austria

Country	Synthetic West Germany Weights	Synthetic Austria Weights	Restricted Synthetic West Germany Weights
West Germany	-	0.33	-
Austria	0.42	-	-
Australia	0	0	0
Belgium	0	0.12	0
Denmark	0	0	0
France	0	0	0
Greece	0	0	0
Italy	0	0	0
Japan	0.16	0.21	0.22
Netherlands	0.09	0.31	0.30
New Zealand	0	0	0
Norway	0	0.03	0
Portugal	0	0	0
Spain	0	0	0
Switzerland	0.11	0	0.09
UK	0	0	0
USA	0.22	0	0.39

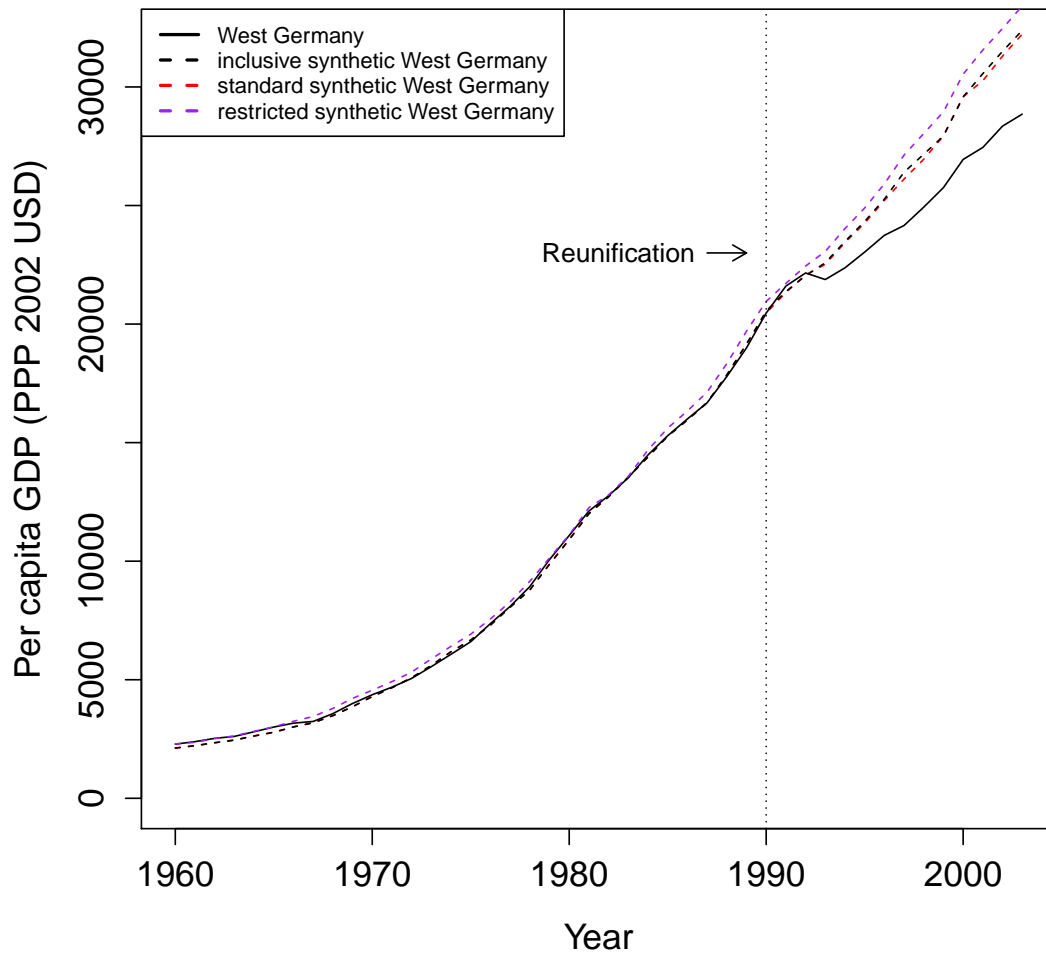


Figure 1: Trends in per capita GDP: West Germany, synthetic West Germany, inclusive synthetic West Germany, and restricted synthetic West Germany

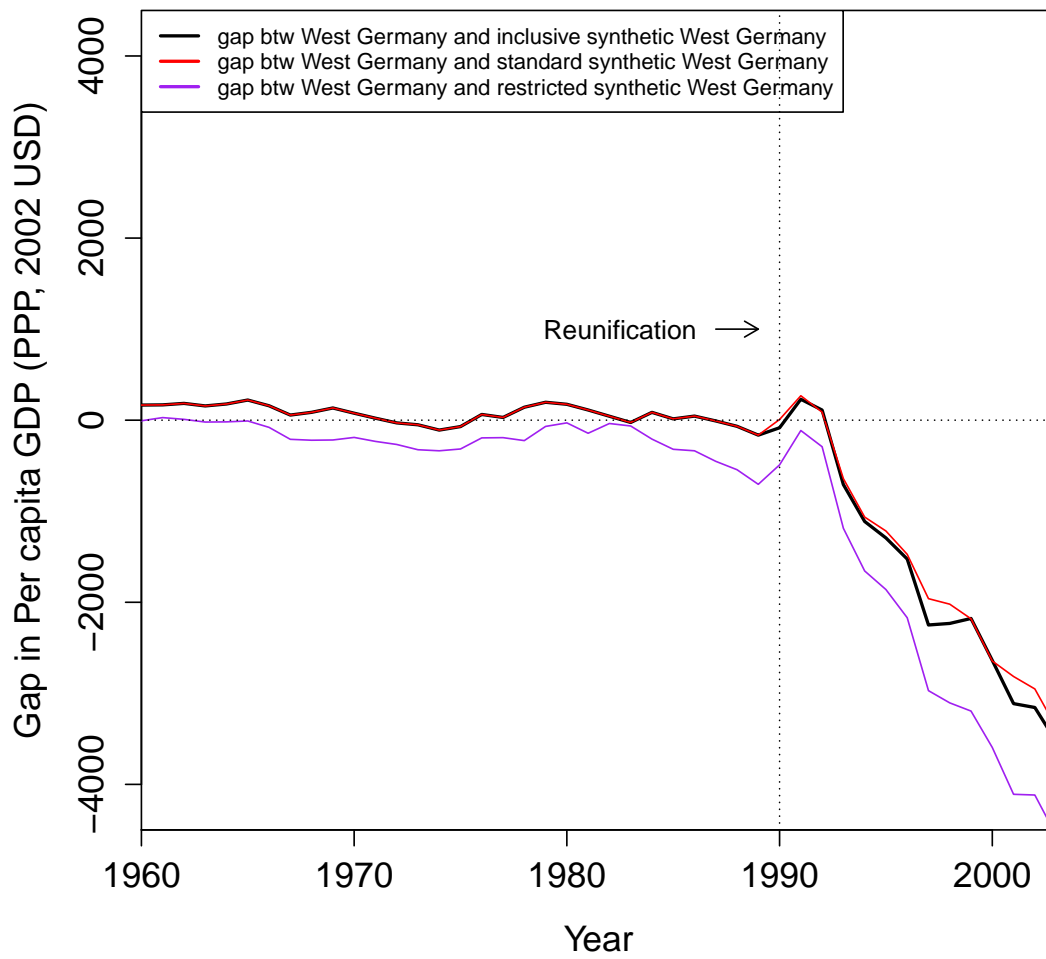


Figure 2: Estimated effects on West Germany

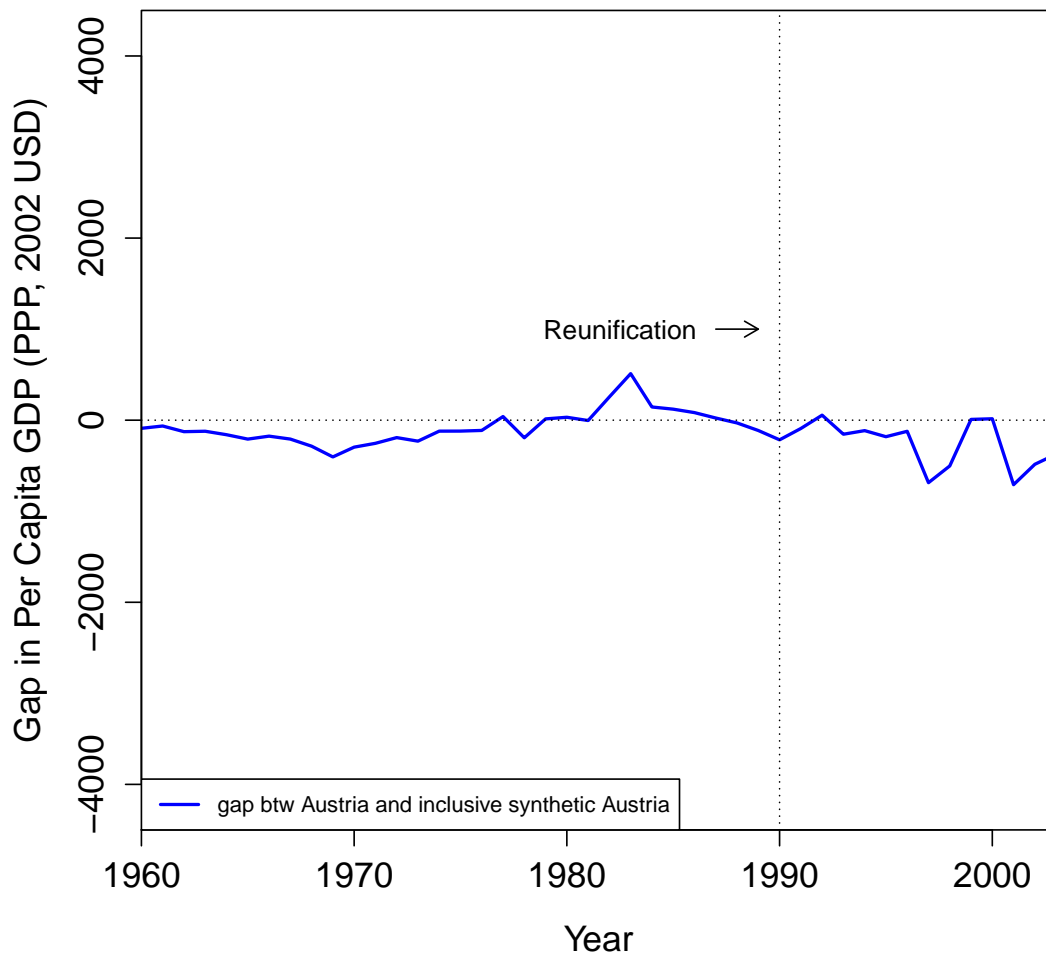


Figure 3: Estimated effects on Austria

Table 2: Treatment Effects on West Germany

	$\theta_t^{SCM}$	$\theta_t^{iSCM}$	$\theta_t^{iSCM} - \theta_t^{SCM}$	$\theta_t^{resSCM}$
1990	7.58	-83.21	-90.79	-490.71
1991	268.31	229.89	-38.42	-113.50
1992	87.90	111.03	23.13	-291.56
1993	-642.23	-707.14	-64.91	-1187.41
1994	-1064.13	-1112.46	-48.33	-1656.09
1995	-1216.99	-1293.31	-76.32	-1860.16
1996	-1473.30	-1524.64	-51.34	-2169.88
1997	-1960.38	-2249.24	-288.86	-2970.69
1998	-2020.74	-2232.20	-211.47	-3104.82
1999	-2181.48	-2177.89	3.59	-3194.84
2000	-2645.30	-2638.79	6.51	-3595.01
2001	-2815.12	-3113.22	-298.10	-4109.12
2002	-2951.69	-3155.55	-203.86	-4116.73
2003	-3372.36	-3529.42	-157.06	-4559.62

Table 3: Treatment Effects on Austria

	$\gamma_t^{SCM}$	$\gamma_t^{iSCM}$	$\gamma_t^{iSCM} - \gamma_t^{SCM}$
1990	-188.12	-215.83	-27.71
1991	-167.88	-91.33	76.55
1992	18.01	54.98	36.97
1993	81.18	-154.29	-235.47
1994	255.56	-114.88	-370.43
1995	249.23	-181.43	-430.66
1996	385.64	-122.04	-507.69
1997	62.32	-686.65	-748.97
1998	240.62	-502.68	-743.29
1999	733.73	8.52	-725.21
2000	894.16	15.48	-878.68
2001	328.05	-708.61	-1036.66
2002	566.15	-484.60	-1050.76
2003	801.89	-373.36	-1175.25

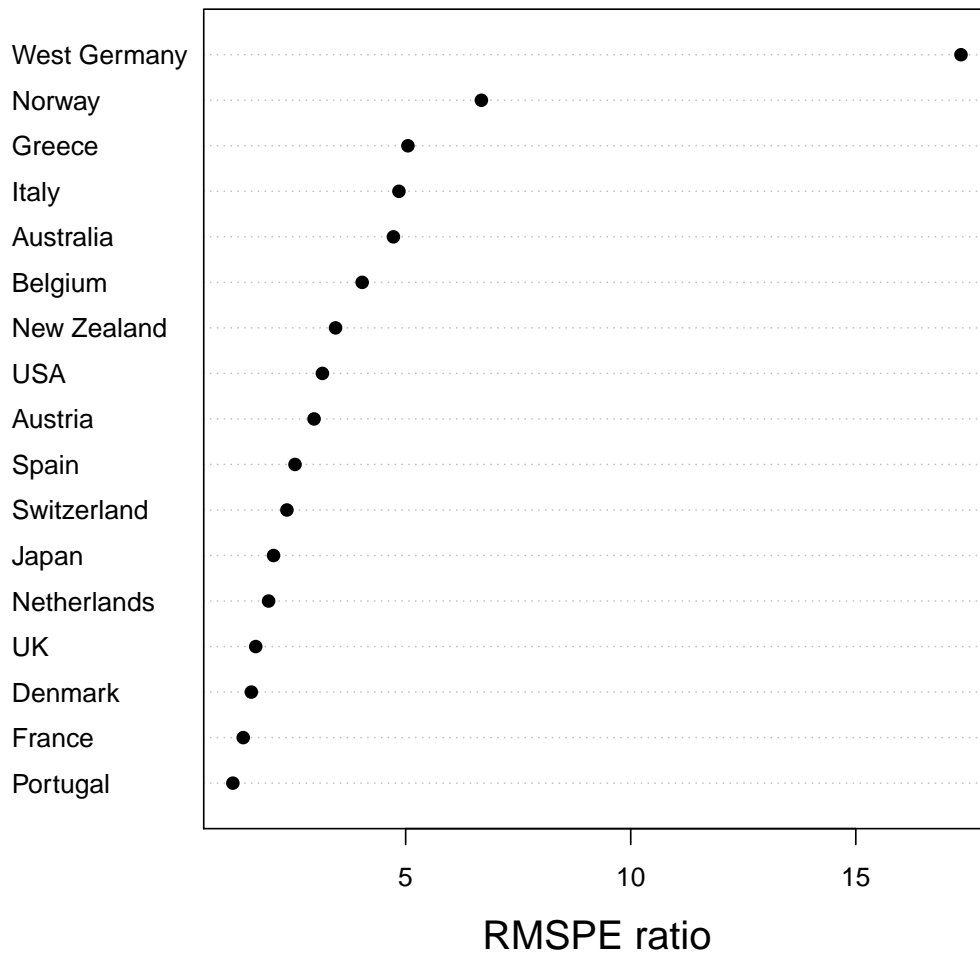


Figure 4: Ratio of post- and pre-reunification RMSPEs: West Germany and control countries

## 6 Conclusion

We introduce iSCM, a modification of the standard SCM, that allows including units potentially affected by an intervention in the donor pool. Our method is useful in applications where it is either important to include other treated units in the donor pool or where some of units are affected indirectly by the intervention (spillover effects). iSCM requires that the assumptions of the standard SCM would be valid in the absence of post-intervention effects as well as the presence of at least one “pure” control unit in the donor pool. A big advantage of iSCM is that it can be easily implemented using the standard synthetic control algorithm or any new estimation method available in the literature. Finally, we illustrate how to use iSCM by estimating the impact of Germany Reunification on GDP per capita, confirming Abadie et al. (2015) expectations about the potential direction of the spillover effect from West Germany to Austria. We find small negative spillover effects to Austria, which would imply an even more negative treatment effect on West Germany.

## Appendix

### A Non-singularity

Let  $\omega_{ij}$  a generic element of  $\Omega$ . We have that

1.  $\omega_{ii} = 1, \forall i = 1, \dots, m$  (the main diagonal elements are all one by definition).
2.  $0 \leq |\omega_{ij}| \leq 1$  (the non-diagonal elements include estimated weights).
3.  $0 \leq \sum_i \omega_{ij} \leq 1$ . (the sum of the weights in a row cannot be bigger than one).
4. If  $|\omega_{ij}| = 1, j \neq i$ , then all the non-diagonal elements on the same row are zero (if one of the weights equals one, all of the others must be zero).

As  $\Omega$  is a square matrix, it is non-singular if, and only if, its determinant is different from zero, which can only be the case if none of the three conditions below are satisfied:

1. Either one of its rows or one of its columns only contains zeros.
2. Either two of its rows or two of its columns are proportional to each other.
3. Either one of its rows or one of its columns is a linear combination of at least two others.

The first and the second conditions are immediately ruled out by the fact that  $\Omega$  always contains ones on its main diagonal and all its other elements are smaller than 1 in absolute value. The third conditions can only occur if either  $\omega_{ij} = \omega_{ji} = -1$ ,  $j \neq i$  or if in every single row we have  $\sum_i \omega_{ij} = 0$ .



## References

- ABADIE, A. (2020): “Using synthetic controls: Feasibility, data requirements, and methodological aspects,” *Journal of Economic Literature*.
- ABADIE, A., A. DIAMOND, AND J. HAINMUELLER (2010): “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program,” *Journal of the American Statistical Association*, 105, 493–505.
- (2015): “Comparative Politics and the Synthetic Control Method,” *American Journal of Political Science*, 59, 495–510.
- ABADIE, A. AND J. GARDEAZABAL (2003): “The Economic Costs of Conflict: A Case Study of the Basque Country,” *American Economic Review*, 93, 113–132.
- ABADIE, A. AND J. L’HOUR (2019): “A Penalized Synthetic Control Estimator for Disaggregated Data,” *working paper*. Available at <https://sites.google.com/site/jeremyhour/research>.
- AMJAD, M., D. SHAH, AND D. SHEN (2018): “Robust synthetic control,” *The Journal of Machine Learning Research*, 19, 802–852.
- ARKHANGELSKY, D., S. ATHEY, D. A. HIRSHBERG, G. W. IMBENS, AND S. WAGER (2019): “Synthetic difference in differences,” Tech. rep., National Bureau of Economic Research.
- ATHEY, S., M. BAYATI, N. DOUDCHENKO, G. IMBENS, AND K. KHOSRAVI (2020): “Matrix completion methods for causal panel data models,” *arXiv:1710.10251v3*.
- ATHEY, S. AND G. W. IMBENS (2017): “The state of applied econometrics: Causality and policy evaluation,” *Journal of Economic Perspectives*, 31, 3–32.
- BEN-MICHAEL, E., A. FELLER, AND J. ROTHSTEIN (2019): “Synthetic Controls and Weighted Event Studies with Staggered Adoption,” *arXiv:1912.03290v1*.
- (2020): “The Augmented Synthetic Control Method,” *arXiv:1811.04170v3*.

- BOTOSARU, I. AND B. FERMAN (2019): “On the role of covariates in the synthetic control method,” *The Econometrics Journal*, 22, 117–130.
- CAO, J. AND C. DOWD (2019): “Estimation and Inference for Synthetic Control Methods with Spillover Effects,” Papers 1902.07343, arXiv.org.
- CERQUA, A. AND G. PELLEGRINI (2017): “Industrial policy evaluation in the presence of spillovers,” *Small Business Economics*, 49, 671–686.
- CHERNOZHUKOV, V., K. WÜTHRICH, AND Y. ZHU (2018): “An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls,” *arXiv:1712.09089v6*.
- DOUDCHENKO, N. AND G. W. IMBENS (2017): “Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis,” *arXiv:1610.07748*.
- FERMAN, B. AND C. PINTO (2019): “Synthetic controls with imperfect pre-treatment fit,” *arXiv preprint arXiv:1911.08521v1*.
- FIRPO, S. AND V. POSSEBOM (2018): “Synthetic control method: Inference, sensitivity analysis and confidence sets,” *Journal of Causal Inference*, 6, 1–26.
- FORASTIERE, L., E. M. AIROLDI, AND F. MEALLI (2016): “Identification and estimation of treatment and interference effects in observational studies on networks,” *arXiv preprint arXiv:1609.06245*.
- GEHLER, M. AND M. GRAF (2018): “Austria, German Unification, and European Integration: A Brief Historical Background,” *working paper-Cold War International History Project*.
- GOBILLON, L. AND T. MAGNAC (2016): “Regional policy evaluation: interactive fixed effects and synthetic controls,” *Review of Economics and Statistics*, 98, 535–551.
- GROSSI, G., P. LATTARULO, M. MARIANI, A. MATTEI, AND Ö. ÖNER (2020): “Synthetic Control Group Methods in the Presence of Interference: The Direct

- and Spillover Effects of Light Rail on Neighborhood Retail Activity,” *arXiv preprint arXiv:2004.05027*.
- HUBER, M. AND A. STEINMAYR (2019): “A framework for separating individual-level treatment effects from spillover effects,” *Journal of Business & Economic Statistics*, 1–15.
- KELLOGG, M., M. MOGSTAD, G. POULIOT, AND A. TORGOVITSKY (2020): “Combining Matching and Synthetic Controls to Trade off Biases from Extrapolation and Interpolation,” *National Bureau of Economic Research*.
- LI, K. T. (2019): “Statistical inference for average treatment effects estimated by synthetic control methods,” *Journal of the American Statistical Association*, 1–16.
- MELLACE, G. AND A. PASQUINI (2019): “Identify More, Observe Less: Mediation Analysis Synthetic Control,” *arXiv e-prints*, arXiv:1909.12073.
- RUBIN, D. B. (1974): “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies,” *Journal of Educational Psychology*, 66, 688–701.
- SCHINASI, G. J., L. LIPSCHITZ, AND D. McDONALD (1990): “Monetary and financial issues in german unification,” *German Unification: Economic Issues. International Monetary Fund, Occasional Paper*, 144–154.
- VAZQUEZ-BARE, G. (2017): “Identification and estimation of spillover effects in randomized experiments,” *arXiv preprint arXiv:1711.02745*.
- XU, Y. (2017): “Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models,” *Political Analysis*, 25, 57–76.